Geotechnical and Geophysical Site Characterization – Huang & Mayne (eds)
© 2008 Taylor & Francis Group, London, ISBN 978-0-415-46936-4

Statistical analysis of geotechnical data

M. Uzielli
Georisk Engineering, Florence, Italy

ABSTRACT: This paper attempts to provide an overview of the main domains of application of statistical-based techniques to geotechnical data. Examples and references to literature contributions are provided.

1 INTRODUCTION

1.1 *Uncertainty, variability and determinism*

The geotechnical engineer processes testing data to obtain parameters for characterization and design. In practice, *information is never sufficient in quantity, nor entirely precise and accurate*. Geomaterials, moreover, are *naturally complex and variable at all scales*, ranging from the microstructure to regional scale. The competent engineer must account for this lack of uniformity and information while parameterizing and modeling the physical world. The *level of explicitness* with which this occurs depends upon the selected approach. In *deterministic approaches*, variability is not addressed explicitly as in *uncertainty-based approaches*.

In the technical literature – and geotechnical engineering is no exception – the terms *variability* and *uncertainty* are often employed interchangeably. Strictly speaking, this is not correct. *Variability* is an observable manifestation of heterogeneity of one or more physical parameters and/or processes. *Uncertainty* pertains to the modeler's state of knowledge and strategy, and reflects the decision to recognize and address the observed variability in a qualitative or quantitative manner.

Deterministic methods lie at the basis of virtually every technological science, and geotechnical engineering is no exception. However, the importance of explicitly modeling and assessing the variability of geotechnical parameters (i.e. quantifying, processing and reporting the associated uncertainty) is increasingly recognized in geotechnical design and characterization. Most evolutionary design codes operate in an uncertainty-based perspective, requiring explicit quantification not only of most suitable values (usually termed 'characteristic' or 'nominal'), but also of the level of uncertainty and confidence in the selection of such values.

The progressive shift towards an uncertainty-based perspective may be motivated by the fact that this may be, on the whole more *convenient* in terms of *safety*, *performance* and *costs*. The explicit parameterization of uncertainty allows to provide more *complete* and *realistic* information regarding the level of risk associated with design. Addressing uncertainty does not *per se* increase the level of safety, but allows the engineer to rationally calibrate his decisions on a desired or required *reliability* or *performance level* of a geotechnical system. Being able to select the performance level and reduce undesired conservatism, in turn, is generally beneficial in the economic sense.

Among the main trade-offs for the positive aspects of uncertainty-based approaches, which hinder a more rapid diffusion among geotechnical practitioners, are the necessity to rely on specialized mathematical techniques and, not infrequently, large computational expense. While ever-increasing computational power is constantly reducing the relevance of the latter, a correct implementation of uncertainty-based techniques requires at least some degree of comprehension on the side of the engineer. The results of uncertainty-based analyses can be used confidently for engineering purposes only if preceded, accompanied and followed by geotechnical expertise and judgment.

1.2 *Rationale and scope of the paper*

Among mathematical disciplines which allow consistent modeling, processing, evaluation and assessment of uncertainty, *statistical theory* (usually employed jointly and iteratively with probability theory) provides a well developed, widely understood and accepted framework. Statistical theory encompasses a broad range of topics. A notable advantage of statistics over other uncertainty-addressing techniques such (e.g. fuzzy logic) is – at present – the vast bulk of statistical software packages which are available.

The main goal of the paper is to provide a wide – though necessarily synthetic – overview of the main domains of applications of statistical analysis of geotechnical data for practical implementation in uncertainty-based characterization and design. This is attempted herein through:

a) Description of the main factors of variability in geotechnical testing data, and definition of the sources of geotechnical uncertainty;
b) Description of the domains of application of statistical analyses for geotechnical purposes;
c) Presentation of selected examples from the geotechnical literature.

Statistical techniques are not addressed from a purely theoretical perspective. Let alone the specialized literature, a number of textbooks are available for a more comprehensive insight into the formal aspects of statistical science in the context of geotechnical engineering: Ang & Tang (1975); Ayyub & McCuen (2003); Baecher & Christian (2003) among others.

It is generally deemed preferable not to 'lose sight of the forest for the trees', i.e. to focus on simpler, more easily applicable methods rather than better-performing techniques which are hardly utilizable in geotechnical practice. The main reason is that there are large, at present unavoidable uncertainties in geotechnical parameters and calculation models which can generally be expected to exceed the potentially existing bias and error resulting from a less refined analysis. In any case, it is attempted to emphasize that *one should never apply even the most simple techniques uncritically.*

This paper focuses on statistical approaches based on the *frequentist* perspective. The rapidly increasing use of Bayesian statistics (which are not addressed here) in geotechnical engineering attests for the advantage of exploiting the dual nature of probability in an integrated, rigorous manner. Interested readers are referred to Ayyub & McCuen (2003), Baecher & Christian (2003) and Christian (2004) for a comparative insight between frequentist and Bayesian statistics in the engineering disciplines.

Many of the procedures which have been developed by geotechnical researchers and practitioners for uncertainty-based analysis make use of statistics. However, they often rely on the synergic integration of statistical methods with other techniques. Hence, the attribute of 'statistical' should not be assessed too rigorously in this paper.

2 VARIABILITY AND UNCERTAINTY IN GEOTECHNICAL TESTING DATA

The variability in geotechnical data can be ascribed to both the soils and the investigators. Soils are natural materials, which are formed and continuously modified by complex processes, as discussed in detail by Hight & Leroueil (2003). The variety and complexity of such processes result in physical heterogeneity and, consequently, in the variability of quantitative parameters. *Inherent soil variability* describes the variation of properties from one spatial location to another inside a soil mass. Inherent soil variability is parameterized by *aleatory uncertainty*, which can be observed at virtually any scale at which properties are measured, and which is not necessarily reduced by increasing the numerosity and quality of data. *Epistemic uncertainty* exists as a consequence of the investigator's invariably limited information and imperfect measurement and modeling capabilities.

2.1 *Shortcomings of total variability analyses*

Uncertainty-based analyses can, in principle, neglect the compound nature of geotechnical uncertainty, and address *total variability* instead. However, there are at least three extremely important reasons for which this approach may not be desirable. The first reason is that analyses performed in the perspective of total variability are strictly site-specific. The strong influence of factors such as the specific testing equipment and personnel, soil type, depositional history and in-situ state parameters on the results of statistical analyses makes it virtually impossible to replicate the overall conditions at any other site (or, possibly, even at the same site at different times or with different equipment). Exporting results of total variability analyses to other sites uncritically generally results in incorrect assessments of uncertainty, and should be avoided as far as possible.

The second main shortcoming of total variability analyses is the overly conservative assessment of uncertainty due to the fact that the important effect of spatial averaging of variability, which is parameterized by a reduction of variance of aleatory uncertainty and which will be described in the paper, cannot be addressed. Total variability analysis thus hinders one of the main goals of uncertainty-based analysis, namely the reduction of excess conservatism in geotechnical characterization and design, or at least a rational assessment of the level of conservatism itself.

Third, a separate assessment of the magnitude of each uncertainty component allows, if desired, specific actions aimed at the reduction of total uncertainty. For instance, if it is seen that the highest contributor to total uncertainty is estimation uncertainty, then it could be decided that more data should be collected. If, on the other hand, measurement uncertainty were deemed too significant, supplementary campaigns making use of more repeatable testing methods could be planned.

2.2 *How useful are literature values?*

Examples and references to literature values of relevant output parameters of statistical analysis are provided in the paper. As the explicit categorization and separation of total uncertainty into aleatory and epistemic components (and sub-components thereof) is not frequently encountered in the geotechnical literature, most literature values should be regarded as deriving from total variability analyses.

On the basis of what has been stated in Section 2.1, while it is conceptually appropriate to view such data as a plausible range of values, the effects of *endogenous* factors (i.e. pertaining to the compositional characteristics of soils) and *exogenous factors* (e.g. related to in-situ conditions, groundwater level and geological history) which may influence the magnitudes of a given geotechnical parameter, should be inspected in the light of geotechnical knowledge if it is of interest to export literature values to other sites.

Hence, literature values should not be exported to other sites without a critical analysis of such endogenous and exogenous factors.

2.3 *Geotechnical uncertainty models*

If it is of interest to bypass the undesirable effects of total variability analyses, a rational assessment of the level of uncertainty in a measured or derived geotechnical parameter can be made by use of an *uncertainty model*. Phoon & Kulhawy (1999b), for instance, proposed an additive model for the total coefficient of variation of a point design parameter obtained from a single measured property using a transformation model:

$$COV_{tot,D}^2 = \frac{\delta}{L} COV_{\omega}^2 + COV_m^2 + COV_{se}^2 + COV_M^2 \quad (1)$$

in which $COV_{tot,D}$ is the total coefficient of variation of the design property; δ/L is an approximation of the *variance reduction* due to *spatial averaging* (see Section 4; δ is the scale of fluctuation of the design property (Section 4.2); L is the spatial extension of interest for design; COV_{ω} is the coefficient of variation of inherent variability of the measured property (Section 4.2); COV_{se} is the coefficient of variation of statistical estimation uncertainty (Section 9.1); COV_m is the coefficient of variation of measurement uncertainty of the measured property (Section 9.2); and COV_M is the coefficient of variation of transformation uncertainty of the transformation model (Section 9.3). The coefficient of variation as a measure of dispersion and uncertainty is defined in Section 3.

In quantitative uncertainty models, the components of uncertainty are usually assumed to be statistically uncorrelated; hence, the absence of correlation terms in Eq. (1). This is only approximately true: for instance, the quantification of variability requires data from measurements of soil properties of interest. Different geotechnical measurement methods, whether performed in the laboratory or in–situ, will generally induce different failure modes in a volume of soil. This usually results in different values of the same measured property. Measured data are consequently related to test-specific failure modes. The type and magnitude of variability cannot thus be regarded as being inherent properties of a soil volume, but are related to the type of measurement technique. Scatter due to small-scale but real variations in a measured soil property is at times mistakenly attributed to measurement error. Given the above, it is extremely difficult to separate the components of geotechnical uncertainty. The hypothesis of uncorrelated uncertainty sources, though approximate, is especially important as it justifies separate treatment of the variability components, thereby allowing the application of the most suitable techniques for each.

Sections 3 through 8 address a number of domains of application of statistical methods for geotechnical characterization and design. The procedures presented therein do not address epistemic uncertainty explicitly. Hence, one should keep in mind, whenever pertinent, that a full quantification of uncertainty requires the adoption of an uncertainty model such as the one in Eq. (1), and the quantitative parameterization of epistemic uncertainties, for instance as described in Section 9.

3 SECOND-MOMENT STATISTICAL ANALYSIS

If a data set pertaining to a specific parameter is addressed statistically, such parameter can be referred to as a *random variable*. The term *sample statistic* refers to any mathematical function of a data sample. An infinite number of sample statistics may be calculated from any given data set. For most engineering purposes, sample statistics are more useful than the comprehensive *frequency distribution* (as given by frequency histograms, for instance). For geotechnical data, given the typically limited size of samples, it is usually sufficient to refer to *second-moment analysis*. In second-moment approaches, the uncertainty in a random variable can be investigated through its first two moments, i.e. the mean (a *central tendency* parameter) and variance (a *dispersion* parameter). Higher-moment statistics such as skewness and kurtosis are not addressed.

The *sample mean*, i.e. the mean of a sample $\xi_1, \ldots, \xi_n$ of a random variable Ξ is given by

$$m_{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i \quad (2)$$

Table 1. Second-moment sample statistics of Szeged soils (adapted from Réthati 1988).

	Soil type	w (%)	w_L (%)	w_P (%)	I_P (%)	I_C	e	S_r	γ (kN/m^3)
Mean	S1	31.1	44.5	24.2	20.6	0.62	0.866	0.917	18.8
	S2	22.3	32.3	19.4	13.0	0.82	0.674	0.871	19.7
	S3	24.5	32.2	20.7	11.5	0.63	0.697	0.902	19.7
	S4	28.3	54.0	24.2	29.9	0.86	0.821	0.901	19.3
	S5	28.6	52.8	25.4	27.4	0.82	0.823	0.905	19.2
COV	S1	0.30	0.39	0.34	0.57	0.49	0.26	0.10	0.07
	S2	0.18	0.13	0.12	0.37	0.37	0.13	0.12	0.04
	S3	0.15	0.13	0.11	0.36	0.47	0.13	0.11	0.04
	S4	0.17	0.21	0.14	0.34	0.20	0.11	0.10	0.03
	S5	0.17	0.27	0.15	0.44	0.32	0.12	0.10	0.03

The *sample variance* of a set of data is the square of the *sample standard deviation* of the set itself. The latter is given by

$$s_\xi = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - m_\xi)^2} \quad (3)$$

The *sample coefficient of variation* is obtained by dividing the sample standard deviation by the sample mean. It provides a concise measure of the relative dispersion of data around the central tendency estimator:

$$COV_\xi = \frac{s_\xi}{m_\xi} \quad (4)$$

The coefficient of variation is frequently used in variability analyses. It is dimensionless and in most cases provides a more physically meaningful measure of dispersion relative to the mean. However, it should be noted that it is a ratio which is very sensitive, for instance in case of small mean values. Harr (1987) provided a 'rule of thumb' by which coefficients of variation below 0.10 are considered to be 'low', between 0.15 and 0.30 'moderate', and greater than 0.30, 'high'.

Though the coefficient of variation is defined as a statistic and, in principle, is to be calculated using Eq. (4), it is paramount to acknowledge the multi-faceted nature of probability: the 'frequentist' perspective relies on *objective* calculation of sample statistics as described previously; the 'degree of belief' perspective requires subjective quantification of input parameters on the basis of experience, belief or judgment. Hence, as shall be discussed in the following, second-moment statistics can be quantified using either perspective.

3.1 *Second-moment analysis of independent random variables*

Parameters such as water content, plasticity index and unit weight are in principle not univocally related to specific in-situ conditions. For instance, they are not affected by the in-situ stress state and, thus, do not always show a trend with depth. These can be defined as *independent random variables* for the purpose of statistical analysis.

As a best-practice literature example of second-moment analysis independent random variables, the results of an investigation on soil physical characteristics (see Réthati 1988) for 5 soil types underlying the town of Szeged (S1: humous-organic silty clay; S2: infusion loess, above groundwater level; S3: infusion loess, below groundwater level; S4: clay; S5: clay-silt) using the results of approximately 11000 tests from 2600 samples are shown. Mean values and COVs of water content w, plasticity index I_P, consistency index I_C, void ratio e, degree of saturation S_r, unit weight γ are reported in Table 1. The subdivision of data by soil type and quality of samples, along with the considerable size of the data samples, ensures a greater significance of results.

3.2 *Second-moment analysis of dependent random variables*

The magnitude and variability of many geotechnical parameters is bound to other parameters. For instance, the undrained shear strength of a cohesive soil generally varies with depth (even if the soil is compositionally homogeneous) due to increasing overburden stress, overconsolidation effects and other in-situ factors. These parameters, when addressed statistically, can be referred to as *dependent random variables.* Such variables generally display *spatial trends* which are due to one or more independent random variables. Trends can be complex and difficult to characterize analytically *a priori* as they are generally a superposition of trends from various independent variables. Recognizing and addressing such trends quantitatively is very important in geotechnical uncertainty-based analysis, because this allows to assess how much of the total spatial variability can be attributed to the independent variables which generate the spatial trend, and how much to the inherent variability of the dependent variable about the trend. *Decomposition* of data from samples of dependent random variables into a deterministic trend and random variation is often achieved using *statistical regression*.

Least-squares regression is widely used to estimate the parameters to be fit to a set of data and to characterize the statistical properties of the estimates. The main outputs of regression analysis are the *regression parameters* which describe the trend analytically, and the parameters which assess the reliability of the output regression model. Such parameters include the *determination coefficient* (commonly denoted R^2 as the square of the *correlation coefficient*), which provides an indication of how much of the variance of the dependent variable can be described by the independent variables which are included in the regression model and how much by the inherent variability, and the *standard error* of the regression model, which equals the standard deviation of the model's errors of prediction, and provides a measure of the uncertainty in the regression model itself. In a best-practice perspective it is very important to report these parameters explicitly along with the regression model (see Figure 2).

Least-squares can be implemented in several versions: *ordinary least-squares* (OLS) relies on the hypothesis of *homoscedasticity* (i.e. constant variance of the residuals) and does not assign weights to data points; generalized least-squares (GLS) relaxes the hypothesis of homoscedasticity in favor of an independent estimate of the variance of the residuals, and allows for weighing of data points as a consequence of the variance model. Today, regression is commonly performed using dedicated software. Far too often, however, regression is approached mechanically and uncritically; the hypotheses underlying specific regression methods are ignored and neglected, leading to incorrect regression models and biased assessment of the uncertainty associated with regression itself. The reader is referred, for instance, to Ayyub & McCuen (2003) for theoretical aspects of regression procedures.

The hypotheses and implications of decomposition are discussed in detail in Uzielli et al. (2006a). It is important to note that there is no univocally 'correct' trend to be identified, but rather a 'most suitable' one. The choice of the trend function must be consistent with the requirements of the mathematical techniques adopted, and must, more importantly, rely on geotechnical expertise. An example of decomposition of undrained shear strength values from consolidated anisotropic undrained triaxial compression tests on Troll marine clays (Uzielli et al. 2006b) is shown in Figure 1.

Another important application of decomposition is in the assignment of 'characteristic' design

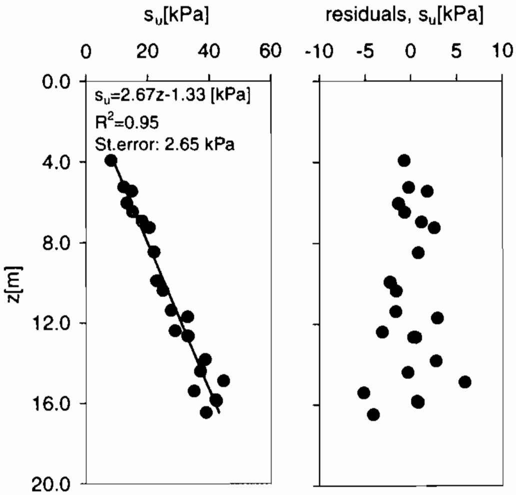

Figure 1. Decomposition of undrained shear strength data from CAUC testing in 2 homogeneous soil units of Troll marine clays: (a) linear trend; (b) residuals of detrending.

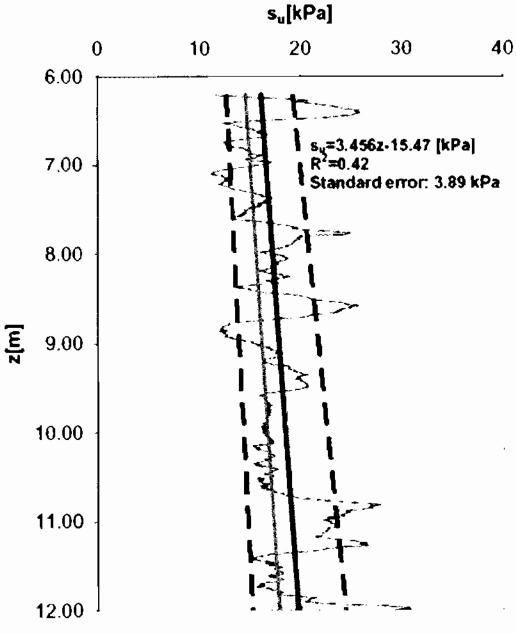

Figure 2. Generalized linear least-squares regression of undrained shear strength calculated from CPTU testing, and subjective design profile (data from Lacasse et al. 2007).

values for dependent random variables. For instance, characteristic values could be taken as the trend function minus a certain number of standard deviations (depending on the degree of desired conservatism). Such approach to the definition of characteristic

Table 2. Multiplicative coefficient for the estimation of the standard deviation of a normally distributed data set with known range (e.g. Snedecor & Cochran 1989).

n	N_n	n	N_n	n	N_n
		11	0.315	30	0.244
2	0.886	12	0.307	50	0.222
3	0.510	13	0.300	75	0.208
4	0.486	14	0.294	100	0.199
5	0.430	15	0.288	150	0.190
6	0.395	16	0.283	200	0.180
7	0.370	17	0.279		
8	0.351	18	0.275		
9	0.337	19	0.271		
10	0.325	20	0.268		

values would reduce the degree of subjectivity in design. Lacasse et al. (2007) compared characteristic design values (previously assigned) of undrained shear strength with the trend obtained from statistical regression on the same data. Figure 2 illustrates an example of such analysis, in which the gray line represents the design values assigned subjectively and the black line is the result of linear generalized least squares regression. It was assessed that there are varying degrees of conservatism in subjectively assigned characteristic values.

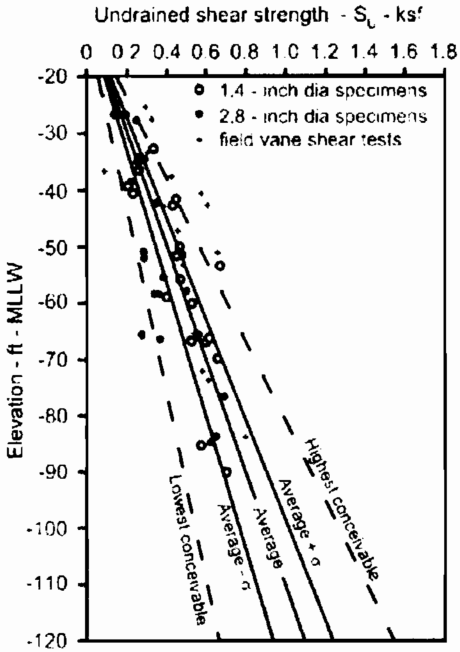

Figure 3. Application of the graphical three-sigma rule for the estimation of the standard deviation of undrained shear strength of San Francisco Bay mud (Duncan 2000).

3.3 *Approximate estimation of sample second-moment statistics*

It may be necessary to assign second-moment parameters in cases in which the complete data set is not available but other sample statistics are known. A number of techniques yielding approximate estimates of sample statistics have been proposed.

For data which can be expected (on the basis of previous knowledge) to be symmetric about its central value, the mean can be estimated as the average of the minimum and maximum values; hence, knowledge of the extreme values would be sufficient. If the range and sample size are known, and if the data can be expected to follow at least approximately a Gaussian (normal) distribution, the standard deviation can be estimated by Eq. (5) (e.g. Snedecor & Cochran 1989), in which the coefficient N_n depends on sample size as shown in Table 2:

$$s_\psi \approx N_n(\psi_{max} - \psi_{min}) \tag{5}$$

3.3.1 *Three-sigma rule*

Dai & Wang (1992) stated that a plausible range of values of a property whose mean value and standard deviation are known can vary between the mean plus or minus three standard deviations. If it is of interest to assign a value to the standard deviation, the statement can be inverted by asserting that the standard deviation can be taken as one sixth of a plausible range of values. The three-sigma rule does not require hypotheses about the distribution of the property of interest even though its origin relies on the normal distribution. In case of spatially ordered data, Duncan (2000) proposed the *graphical three-sigma rule method*, by which a spatially variable standard deviation can be estimated. To apply the method, it is sufficient to select (subjectively or on the basis of regression) a best-fit line through the data. Subsequently, the *minimum* and *maximum conceivable bounds lines* should be traced symmetrically to the average line. The *standard deviation lines* can then be identified as those ranging one-third of the distance between the best-fit line and the minimum and maximum lines. An application of the graphical three-sigma rule (Duncan 2000) is shown in Figure 3. While the three-sigma rule is simple to implement and allows exclusion of outliers, Baecher & Christian (2003) opined that the method may be significantly unconservative as persons will intuitively assign ranges which are excessively small, thus underestimating the standard deviation. Moreover, for the method to be applied with confidence, the expected distribution of the property should at least be symmetric around the mean, which is not verified in many cases.

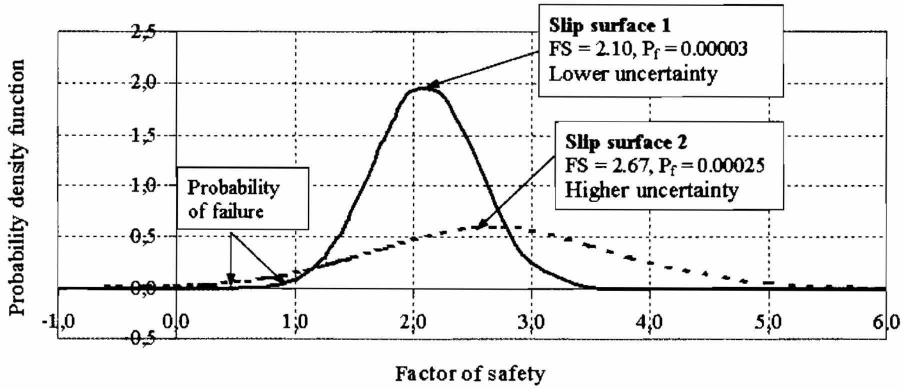

Figure 4. Comparative deterministic and probabilistic assessment of stability for 2 slip surfaces (Nadim & Lacasse 1999).

An exemplifying case which shows the importance of probabilistic approaches even at the simplest second-moment level is provided by Nadim & Lacasse (1999), who performed an undrained slope stability analysis on a slope consisting of 2 clay layers under static and seismic loading using Bishop's method. Analyses were performed both deterministically (i.e. evaluating the factor of safety) and probabilistically (i.e. estimating the probability of failure) using the First-Order Reliability Method (FORM). In the latter case, undrained strength, unit weight and model uncertainty (the latter consisting of a 'mean' model bias term and a COV of model dispersion) were modeled in the second-moment sense. Results showed that, due to the effect of spatial averaging phenomena, the critical surface with the lowest deterministic safety factor was not the critical surface with the highest probability of failure (Figure 4). This example attests for the importance of the additional information which can be provided by probabilistic approaches in comparison with methods which are solely deterministic.

4 SPATIAL VARIABILITY ANALYSIS

Second-moment statistics alone are unable to describe the spatial variation of soil properties, whether measured in the laboratory or in-situ. Two sets of measurements may have similar second-moment statistics and frequency distributions, but could display substantial differences in spatial distribution. Figure 5 provides a comparison of two-dimensional spatial distribution of a generic parameter ξ having similar second-moment statistics and distributions (i.e. histograms), but different degrees of spatial correlation: weak correlation (top right) and strong correlation (bottom right).

Knowledge of the spatial behavior of soil properties is thus very important in uncertainty-based geotechnical analysis and design, at least because: (a) geotechnical design is based on site characterization, which objective is the description of the spatial variation of compositional and mechanical parameters of soils; (b) the values of many geotechnical parameters depend on in-situ state factors (e.g. stress level, overconsolidation ratio, etc.) which are related to spatial location; (c) quantification of the magnitude of spatial variability can contribute significantly to reducing the degree of conservatism in design by allowing the inclusion of the spatial averaging effect in geotechnical uncertainty models.

The goal of spatial variability analyses is to assess quantitatively how much and in which way a given property changes along pre-determined spatial orientations. The vast majority of spatial variability analyses rely on the calculation of spatial statistical properties of data sets. Higher-level analyses require the results of lower-level analyses, and allow consideration of additional effects and parameters which improve the quality of the results. The trade-off for the simplicity of lower-level analyses lies in the limited generality of results as well as in the imperfect modeling of the behavior of geotechnical systems. A detailed insight into such techniques, along with a comparative example of different levels of analysis for slope stability evaluation, is provided by Uzielli et al. (2006a).

4.1 *Spatial averaging of variability*

An implicit manifestation of spatial correlation which is commonly encountered in geotechnical engineering practice is that the representative value of any soil property depends on the volume concerned in the problem to be solved. With reference to a given soil unit and to a specific problem, the geotechnical engineer is trained to define the design values of relevant parameters on the basis of the magnitude of the volume of soil governing the design. Any laboratory or in-situ geotechnical measurement includes some degree of spatial averaging in practice, as tests

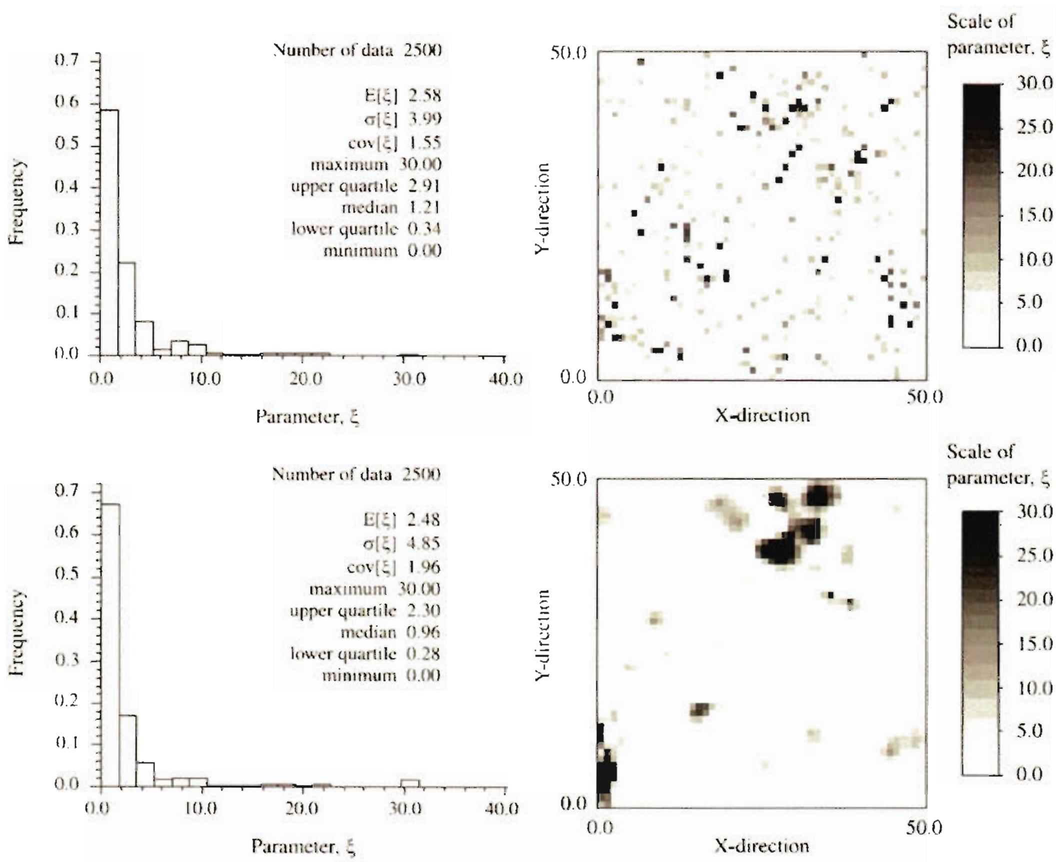

Figure 5. Comparative representation of spatial data with similar statistical distributions (top and bottom right) but different magnitudes of spatial correlation: weak correlation (top right) and strong correlation (bottom right) (from El-Ramly et al. 2002).

are never indicative of point properties, but, rather, are used to represent volumes of soil. The spatial averaging effect results in a reduction of the effect of spatial variability (and, hence, of excessive conservatism) on the computed performance because the variability (in statistical terms, variance) is averaged over a volume, and only the averaged contribution to the uncertainty is of interest as it is representative of the 'real' physical behavior.

4.2 *Random field modeling of spatial variability*

Among the available techniques to investigate spatial variability, *random field theory* has been most frequently referred to in the geotechnical literature. A *random field* is essentially a set of values which are associated to a one- or multi-dimensional space. Values in a random field are usually spatially correlated, meaning that spatially adjacent values can be expected (on the basis of statistical and probability theory) not to differ as much as values that are further apart. It can be described (in the second-moment sense) by its mean, standard deviation (or coefficient of variation) and scale of fluctuation, as well as by an autocorrelation function, which describes the way in which a given property is correlated in space.

The scale of fluctuation is a concise indicator of the spatial extension of correlation. Within separation distances smaller than the scale of fluctuation, the deviations from the trend function are expected to show significant correlation. When the separation distance between two sample points exceeds the scale of fluctuation, it can be assumed that little correlation exists between the fluctuations in the measurements. Though a dependence on soil type has been noted (Uzielli et al. 2005a; 2005b), the scale of fluctuation is not an inherent property of a soil parameter and, when it exists, can be estimated using a variety of methods such as autocorrelation model fitting, calculation of the variance reduction function and semivariogram fitting. These methods, which basically rely on the statistical investigation of spatial correlation properties of the random field, are described in detail in Uzielli et al. (2006a). The scale of fluctuation is also useful to approximate the variance reduction function as shown in Eq (1).

In random field analysis of dependent random variables, it is not conceptually correct to estimate the coefficient of variation on the raw data set as the presence of a spatial trend, (for instance due to effective overburden stress or overconsolidation) would introduce bias in such estimate. Phoon & Kulhawy (1999a) defined the *coefficient of variation of inherent variability* more rigorously as the ratio of the standard deviation of the residuals of data detrending to the mean value of the spatial trend. Very few estimates of the 'real' coefficient of variation of inherent variability are available in the literature; these are reported in Uzielli et al. (2006a). Calculation of the COV as described above provides a more realistic assessment of the dispersion of values around a spatial trend as discussed in Section 3.2.

Uzielli et al. (2006a) provided an extensive list of references to geotechnical spatial variability analyses. Recent contributions include Jaksa (2006), Cherubini et al. (2006), Chiasson & Wang (2006) and Uzielli et al. (2006b).

4.3 *Advanced analysis of geotechnical systems*

The description of a random field in the second-moment sense through a mean, standard deviation, a scale of fluctuation and a spatial correlation function is useful to characterize a spatially variable soil property. However, some possible limitations in this approach should be recognized. For instance, if spatial variability of soil properties is included in an engineering model, stresses and/or displacements which would not appear in the homogeneous case (i.e. in which variability is not addressed) could be present. Random field theory alone is unable to model the influence of spatial variability in the *behavior* of geotechnical systems.

A number of studies have focused, in recent years, on the combined utilization of random fields, non-linear finite element analysis and Monte Carlo simulation for investigating the behavior and reliability of geotechnical systems when the variability of soil properties which are relevant to the main presumable failure mechanisms is considered. Some important general observations can be made on the basis of their results. First, when soils are modelled as spatially variable, the modelled failure mechanisms are quite different – and significantly more complex – than in the case of deterministic soil properties. Second, there generally exists a *critical correlation distance* which corresponds to a minimum-reliability state. Third, phenomena governed by highly non-linear constitutive laws are affected the most by spatial variations in soil properties.

Popescu et al. (2005), for instance, investigated the differential settlements and bearing capacity of a rigid strip foundation on an overconsolidated clay layer. The undrained strength of the clay was modeled as a non-Gaussian random field. The deformation modulus was assumed to be perfectly correlated to undrained shear strength. Ranges for the probabilistic descriptors of the random field were assumed from the literature. Uniform and differential settlements were computed using non-linear finite elements in a Monte Carlo simulation framework. Figure 6a shows the contours of maximum shear strain for a uniform soil deposit with undrained strength of 100 kPa and for a prescribed normalized vertical displacement at center of foundation $\delta/B = 0.1$. In this case the failure mechanism is symmetric and well-defined. The results of the analyses indicated that different sample realizations of soil properties corresponded to fundamentally different failure surfaces. Figure 6b shows an example of a sample realization in which the spatial distribution of undrained strength is not symmetric with respect to the foundation. Hence, as could be expected, the configuration at failure, shown in Figure 6c, involves a rotation as well as a vertical settlement. The repeated finite-element analysis allows appreciation of compound kinematisms (settlements and rotations) of the footings, which could not be inferred from bearing capacity calculations involving non-variable parameters. It was observed that failure surfaces are not mere variations around the deterministic failure surface; thus, no 'average' failure mechanisms could be identified. Another notable result was the observed significant reduction in the values of bearing capacity spatially in the heterogeneous case in comparison with the deterministic model. Figure 6d shows that the normalized pressure which causes a given level of normalized settlement is always higher in the deterministic case. A number of other studies implementing integrated, statistics-based methods for the advanced analysis of geotechnical systems are presented and discussed in Uzielli et al. (2006a).

5 IDENTIFICATION OF PHYSICALLY HOMOGENEOUS SOIL VOLUMES

Second-moment statistics of geotechnical properties available in the literature are generally associated to soil types (e.g. 'clayey soils'). The identification of physically homogeneous units from tested soil volumes is an essential prerequisite for meaningful soil characterization and, consequently, for reliable design. From a qualitative point of view, if analyses are performed on data sets from, say, interbedded stratigraphic units, their results can be expected to be less confidently related to a specific soil type.

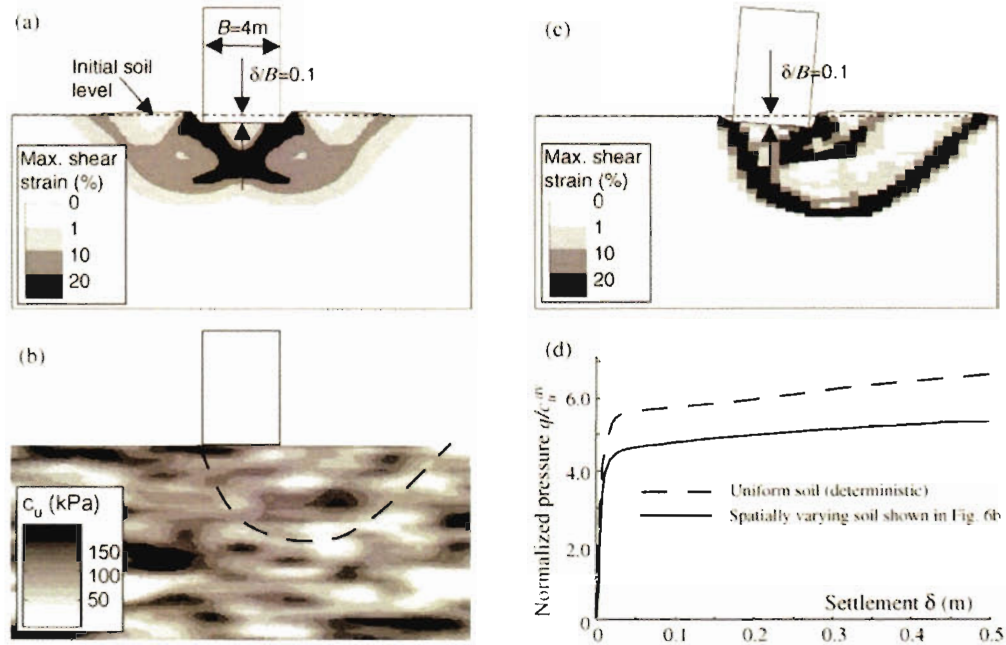

Figure 6. Selected results of investigation on homogeneous and spatially random foundation soil (Popescu et al. 2005).

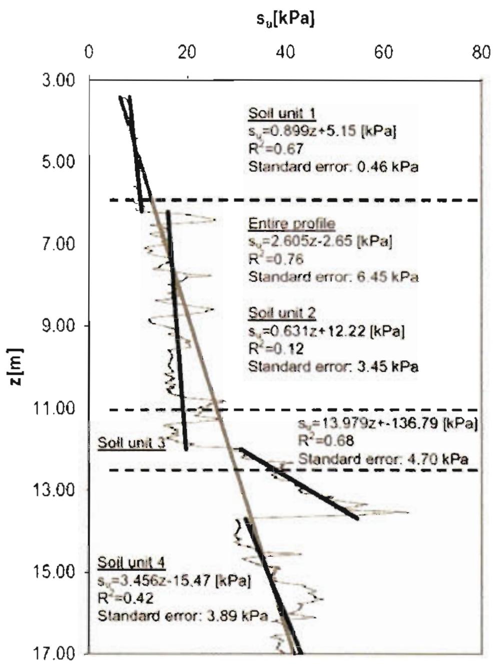

Figure 7. Single-unit vs. multi-unit linear trends in view of the assignment of design values of undrained shear strength (data from Lacasse et al. 2007).

Correspondingly, from a quantitative perspective, resulting statistics (and derived design parameters) can be expected to be less precise and accurate, and inferences made for engineering purposes less refined. Figure 7, for instance, illustrates the effect of subdividing the data from a CPTU sounding into homogeneous units on linear detrending of undrained shear strength values. The gray line represents the linear trend which is obtained by generalized least squares regression on the entire profile; the black lines correspond to the trends in each individual homogeneous soil layer. It may be seen that the unit-level approach provides a set of linear trends which are significantly more respondent to the undrained strength profile. The fact that the determination coefficient is highest for the entire-profile approach indicates that the linear regression model explain a higher quota of the variation of s_u with depth that the unit-level trends, for which inherent variability plays a more relevant role. Hence, design values based on such approach (for instance, taken as the linear trend minus one standard deviation) would be, depending on the depth, beneficially less over-conservative or less under-conservative.

Examples of both subjective and objective assessment of physical homogeneity are available in the geotechnical literature. A purely subjective assessment relying uniquely on expert judgment may not provide optimum and repeatable results. A purely objective assessment based exclusively on numerical criteria

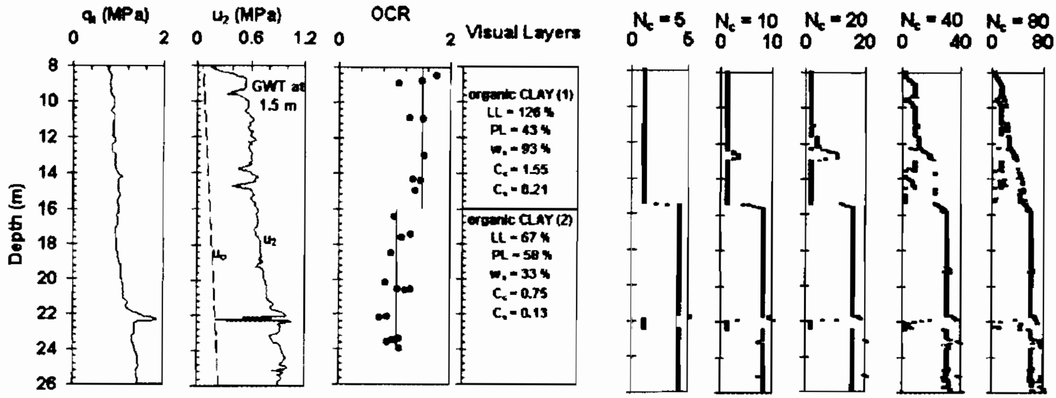

Figure 8. Identification of stratigraphic boundaries from cluster analysis – comparative configurations for variable number of clusters on same profile of corrected cone tip resistance (Hegazy & Mayne 2002).

is not feasible in practice, as at least the preliminary selection and assessment of data quality and the definition of relevant parameters require geotechnical expertise. Hence, in practice, homogeneity assessment is always both subjective and objective to some degree.

In geotechnical analyses, physical homogeneity can be assessed in terms of either soil *composition* or soil *behavior*. These do not display a one-to-one correspondence, as soils which may be homogeneous in terms of composition may not be so in terms of mechanical behavior due to inherent properties of the geomaterials, in-situ state and imperfections in classification systems. Uzielli & Mayne (2008) presented an example of CPTU-based classification in which the Robertson (1990) charts provide heterogeneous categorization for a highly plastic calcareous clay due to strong cementation. Hence, it is essential to define the criteria adopted for physical homogeneity assessment and to report them when presenting results. These aspects are discussed in greater detail in Uzielli et al. (2006a).

Statistical-based procedures have been proposed in the geotechnical literature with the aim of providing increasingly repeatable and objective criteria for the assessment of soil homogeneity. Two examples are illustrated in the following. The aim is not to replace or downplay the importance of judgment; rather, to provide a more solid framework for the rational application of geotechnical expertise.

Hegazy & Mayne (2002) proposed a method for soil stratigraphy delineation by cluster analysis of piezocone data. The objectives of cluster analysis applied to CPT data were essentially: (a) objectively define similar groups in a soil profile; (b) delineate layer boundaries; and (c) allocate the lenses and outliers within sub-layers. Cluster analysis, allows identification of stratigraphic boundaries at various levels of resolution: if a configuration with a low number of clusters is referred to (e.g. 5 cluster-configuration in Figure 8), the main boundaries are identified; with increasing number of clusters, greater resolution is achieved, by which less pronounced discontinuities are captured. Other applications of statistical clustering to CPT data include Facciorusso & Uzielli (2004).

Uzielli (2004) proposed a statistical moving window procedure to identify physically homogeneous soil units. Each moving window is made up of two semi-windows of equal height above and below a centre point. A suitable range of moving window height could be established as 0.75–1.50 m on the basis of soil type, test measurement interval and the type of soil failure (bearing capacity, shear) induced by the testing method. At each center point, user-established statistics (e.g. range, COV) are calculated for data lying in the interval corresponding to the upper and lower limits of the moving window. Homogeneous soil units are essentially identified by delineating soundings into sections where the values of the aforementioned statistics do not exceed preset thresholds. The magnitudes of such thresholds represent the degree of required homogeneity. A minimum width for homogeneous units can be preset by the user to ensure statistical numerosity of samples. Figure 9 shows the application of the method to cone penetration testing data from the heavily interbedded stratigraphy of Venice lagoon soils (Uzielli et al. 2008). Applications of the moving window procedure also include Uzielli et al. (2004; 2005a; 2005b).

The performance of the homogeneous soil unit identification methods should always be assessed critically using geotechnical knowledge and statistical procedures. For instance, a variety of soil classification charts based on in-situ tests are available in the geotechnical literature; these provide an effective means for subjective assessment of the homogeneity of a data set. If data points from a soil unit plot as a well-defined cluster (with possible limited outliers, if

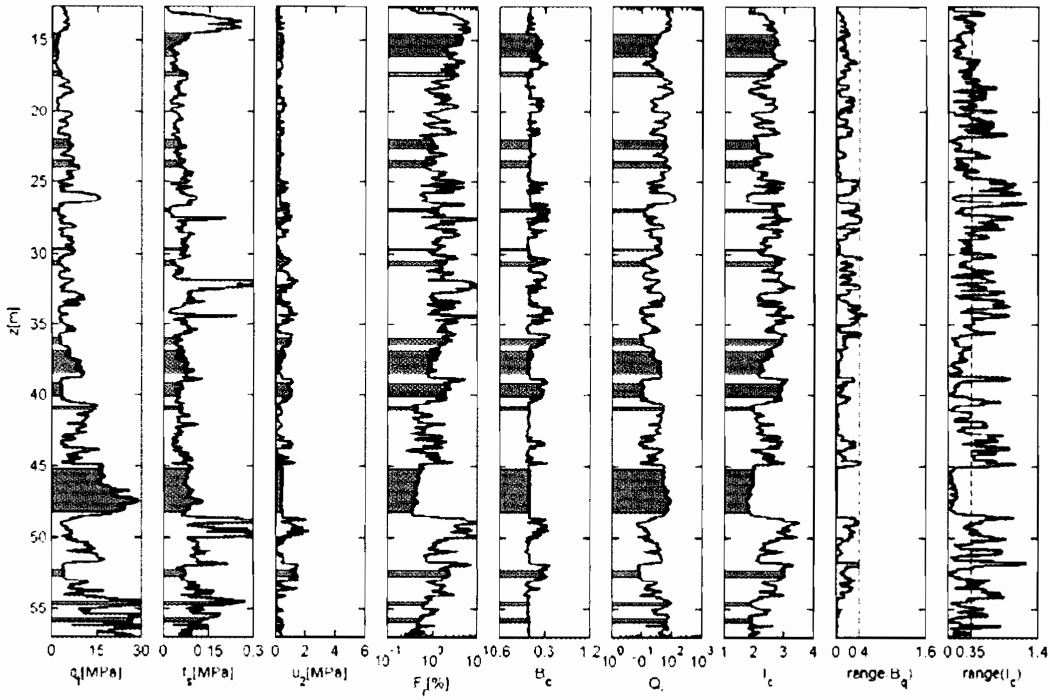

Figure 9. Example identification of physically homogeneous soil layers (in gray) from CPTU test data in Venice lagoon soils (Uzielli et al. 2008).

tolerated by the user) in such charts, homogeneity can be expected. Figure 10 illustrates an example application to dilatometer data using the chart by Marchetti & Crapps (1981), which is useful for DMT-based classification. Examples for CPT data are available in Uzielli (2004) and Uzielli et al. (2004; 2005a; 2005b). The performance of the moving window procedure can also be assessed objectively. Uzielli et al. (2005a) employed a statistical check to assess the performance of the moving window statistical methodology illustrated previously by calculating the coefficient of variation of the soil behavior classification index I_c (which is an efficient descriptor of soil behavior) in each identified homogeneous soil unit.

6 GEOSTATISTICAL ESTIMATION OF SOIL PROPERTIES

In geotechnical site characterization or for design purposes, it is often desirable or necessary to acquire a multi-dimensional spatial map of parameters of interest. As measurements are almost invariably 'too few' or 'not where they would be most useful', it becomes necessary to estimate soil properties at specific locations where observations are not available. The term *kriging* encloses a set of geostatistical techniques which allow such inferential estimation. Kriging techniques are essentially weighted, moving average statistical interpolation procedures which minimize the estimated variance of interpolated values with the weighted averages of their neighbors. The input information required for kriging includes: available data values and their spatial measurement locations; information regarding the spatial correlation structure of the soil property of interest; and the spatial locations of target points, where estimates of the soil property are desired. The weighting factors and the variance are computed using the information regarding the spatial correlation structure of the available data. Since spatial correlation is related to distance, the weights depend on the spatial location of the points of interest for estimation. Formal aspects of kriging can be found, for instance, in Journel & Huijbregts (1978), Davis (1986) and Carr (1995). Two examples from the literature are summarized in the following.

Figure 11a (Lacasse & Nadim 1996) shows the locations of cone penetration test soundings in the neighborhood of a circular-shaped shallow foundation which was to be designed; Figure 11b reports the superimposed profiles of cone resistance in the soundings. It was of interest to estimate the values of cone resistance in other spatial locations under the design location of the foundation itself. Figure 12

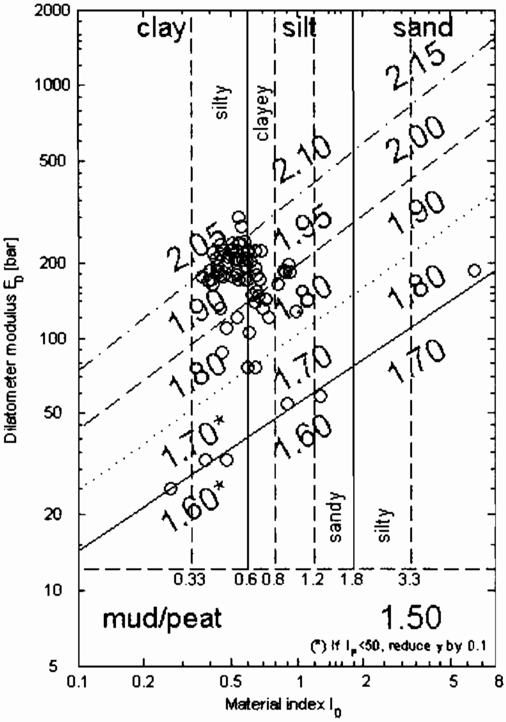

Figure 10. Assessment of the statistical moving window method: plotting data from a physically homogeneous soil unit on Marchetti and Crapps' (1981) I_D-E_D chart.

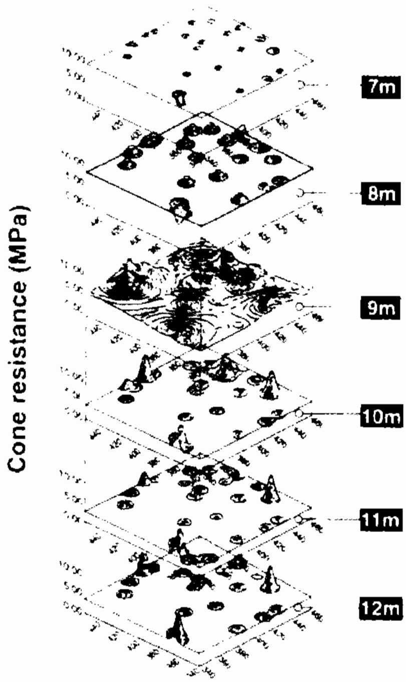

Figure 12. Contours of cone penetration resistance at each meter as obtained by geostatistical kriging (Lacasse & Nadim 1996).

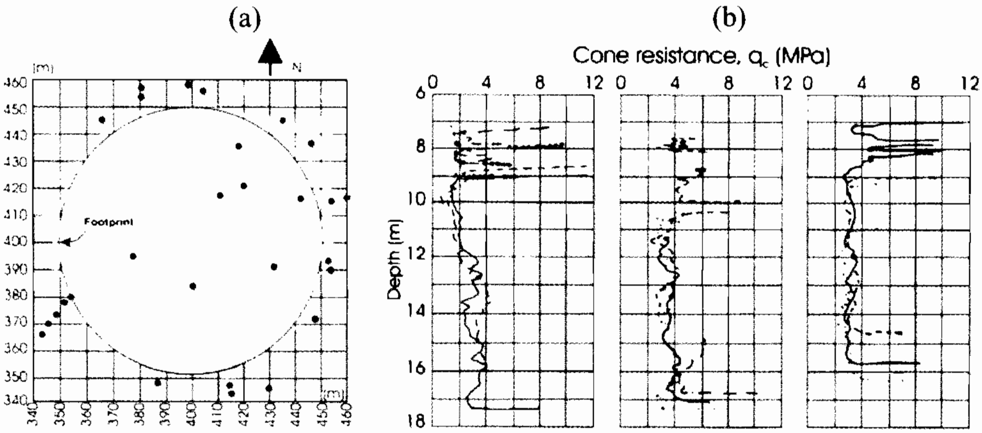

Figure 11. Input data for geostatistical analyses: (a) locations of cone penetration test soundings; (b) superimposed profiles of cone resistance (Lacasse & Nadim 1996).

presents the contours of cone penetration resistance at each meter as obtained by kriging. The 3D graphic representation provides improved insight into the possible spatial variation of cone resistance and the most likely values beneath the foundation. The results of the analysis enabled designers to determine more reliably the position of a clay layer and to use higher shear strength in design, thus reducing conservatism.

Baise et al. (2006) presented an integrated statistical method for characterizing geologic deposits for liquefaction potential using sample-based liquefaction probability values. The method consists of three steps, namely: (1) statistical characterization of samples; (2) evaluation of the spatial correlation; and (3) estimation of the distribution of high liquefaction probability values. In presence of spatial correlation, kriging was used to evaluate the spatial clustering of high liquefaction probability in view of a regional liquefaction potential characterization. An example output is shown in Figure 13.

7 CORRELATION, STATISTICAL DEPENDENCE AND STATIONARITY ASSESSMENT

When dealing with more than one random variable, uncertainties in one may be associated with uncertainties in another, i.e. the uncertainties in the two variables may not be independent. Such dependency (which may be very hard to identify and estimate) can be critical to obtaining proper numerical results in engineering applications making use of well-known probabilistic techniques such as First-Order Second-Moment approximation (FOSM) and First-Order Reliability Method (FORM). Neglecting correlation may result in overdesign or underdesign, depending on the nature of the correlation.

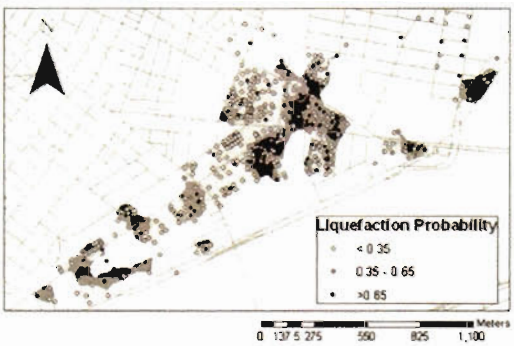

Figure 13. Interpolated map of liquefaction probability in an artificial fill obtained by kriging (Baise et al. 2006).

There are two general approaches to testing for statistical dependence: *parametric* and *nonparametric*. Parametric approaches require the formulation of hypotheses regarding the nature and distribution of the data set under investigation. Non-parametric approaches, on the contrary, do not make such basic assumptions. Consequently, the latter are more widely applicable than parametric tests which often require normality in the data. While more widely applicable, the trade-off is that non-parametric tests are less powerful than parametric tests.

The most common measure of statistical dependence among random variables is *Pearson's linear correlation coefficient*. This measures the degree to which one uncertain quantity varies *linearly* with another. The correlation coefficient is non-dimensional, and varies in the range $[-1, +1]$. A higher bound implies a strict linear relation of positive slope (e.g. Figure 14), while the lower bound attests for a strict linear relation of negative slope. The higher the magnitude, the more closely the data fall on a straight line. Uzielli et al. (2006a) presented several literature examples of Pearson's correlation coefficient between different soil properties at selected sites. The clause of linearity should not be overlooked when interpreting the meaning of correlation: two uncertain quantities may be strongly related to one another, but the resulting correlation coefficient may have negligible value if the relationship is non-linear (e.g. Figure 14).

To overcome the sensitivity of correlation parameters on linearity, it is necessary to refer to procedures which are capable of measuring the strength of the dependence between two variables regardless of the type of relation. A number of statistical tests are available to perform such investigation. *Kendall's tau test* (e.g. Daniel 1990) involves the calculation of the test statistic, τ_{ken}, which measures the degree of concordance (or discordance) between two random variables, i.e. the strength of the trend by which one variable increases (or decreases) if the other variable increases (or decreases). Hence, linearity of the relationship is no longer a constraint. The values of τ_{ken} range from -1 to $+1$, indicating, respectively, perfect negative and positive correlation; values close to zero indicate weak dependence. Critical values of τ_{ken} for rejecting the

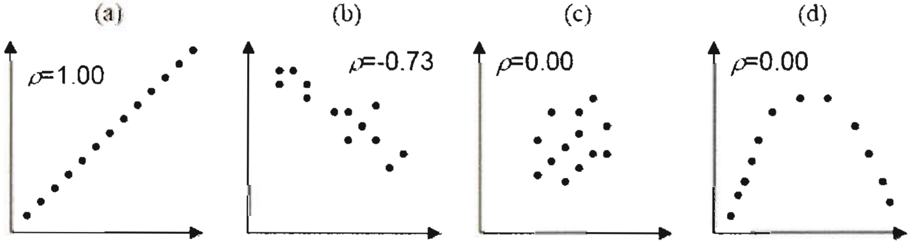

Figure 14. Examples of correlation coefficient between two variables.

null hypothesis of independence are available in tabulated form (e.g. Daniel 1990) or, for larger data sets, assuming that the standardized τ_{ken} statistic is normally distributed, using standard normal distribution tables for a desired confidence level. Kendall's tau test can also be used to establish, when in doubt, whether a random variable should be addressed as being dependent or independent, i.e. to check whether there is a significant spatial trend which would result in dependency. Application of Kendall's test to geotechnical data can be found in Jaksa (1995) and Uzielli et al. (2004).

7.1 *Weak stationarity*

Stationarity is often an important prerequisite for spatial variability analyses of soil properties because many statistical procedures employed therein are based on the assumption that data samples consist of at least *weakly stationary* observations. A process is said to be stationary in the second-order sense (or weakly stationary) if: a) its mean is constant (i.e. there are no trends in the data); b) its variance is constant; and c) the correlation between the values at any two points at distinct spatial locations, depends only on the interval in the spatial distance between the points, and not on the specific spatial location of the points themselves. In the case of laboratory or in-situ geotechnical testing, stationarity can generally only be identified in a weak or second-order sense because of the limitations in sample size. Weak stationarity is discussed in greater detail in Uzielli et al. (2006a).

As statistical independence implies stationarity (the converse is not true), procedures designed for the assessment of the former could be (and have often been) used for assessing the latter. However, the matter should be handled with caution. Classical parametric and non-parametric statistical independence tests such as Kendall's tau test are based on the assumption of spatially uncorrelated input data. This assumption is antithetical to the fundamental premise in geotechnical spatial variability analysis, namely the existence of a correlation structure in spatially varying soil properties. Due to the resulting bias in autocorrelation coefficients in the correlated case, the application of such tests may result in unconservative assessments, i.e. non-stationary sets may be erroneously classified as weakly stationary.

The MBSR test, proposed by Uzielli (2004) as an evolution of a procedure proposed by Phoon et al. (2003), explicitly includes the hypothesis of correlation in data, and allows overcoming of the aforementioned bias. The test consists essentially in the comparison between the maximum value B_{max} of a 'Bartlett statistic', calculated from a moving sampling window, and a critical value B_{crit}. The latter is calculated from *profile factors* which depend on: the degree of spatial correlation of a given soil property of interest (parameterized by the *scale of fluctuation*); the measurement interval of available data; the spatial length of the soil record of length; and the spatial extension of the statistical sampling window assigned by the investigator. Critical values at 5% level of significance for several autocorrelation models, among which single exponential, cosine exponential, second-order Markov and squared exponential, were provided by Phoon et al. (2003). The null hypothesis of stationarity in the variance is rejected at 5% level of significance if $B_{\max} > B_{\text{crit}}$. The MBSR method is described in detail in Uzielli et al. (2006a). Example applications of the method are provided in Uzielli et al. (2004; 2005a; 2005b).

8 SOIL CLASSIFICATION

Soil classification based on laboratory and in-situ testing is invariably affected by considerable uncertainty. In most general terms, the main sources of uncertainty are testing uncertainty (including the effects of sample disturbance) and (especially in the case of in-situ testing, where samples are seldom retrieved) the uncertainty due to the lack of direct information regarding compositional properties of penetrated soils. The two sources of uncertainty can be assumed to be independent, and can be investigated separately. The effect of testing uncertainty is illustrated using an example in Section 9.2. The lack of perfect correlation between mechanical characteristics of soils (which depend on the mode of failure induced by a specific testing technique) and their composition is the main contributor to uncertainty in classification. Soil classification from CPT testing, for instance, relies on the mechanical response of soils to cone intrusion at least in terms of bearing capacity (for cone resistance) and shear failure (for sleeve friction). Due to the inherent complexity of geomaterials and to the variability of in-situ conditions from one testing campaign to another, mechanical properties cannot be related univocally to compositional characteristics. Hence, sets of measurements yielding the same numerical values may pertain to soil volumes which are compositionally different.

Quantitative uncertainty estimation in soil classification is deemed more useful to the geotechnical engineer than qualitative uncertainty assessment, as the most widely used classification methods based on in-situ tests (e.g. Robertson's 1990 chart-based system) operate in a quantitative sense. Moreover, the results of soil classification are used in several engineering applications, for instance liquefaction susceptibility evaluation (e.g. Robertson & Wride 1998).

Zhang & Tumay (2003 and previous) proposed two uncertainty-based classification approaches (one *probabilistic*, based on probability density functions, and one *possibilistic*, based on fuzzy sets) making use of CPT data. The approaches aim to quantify

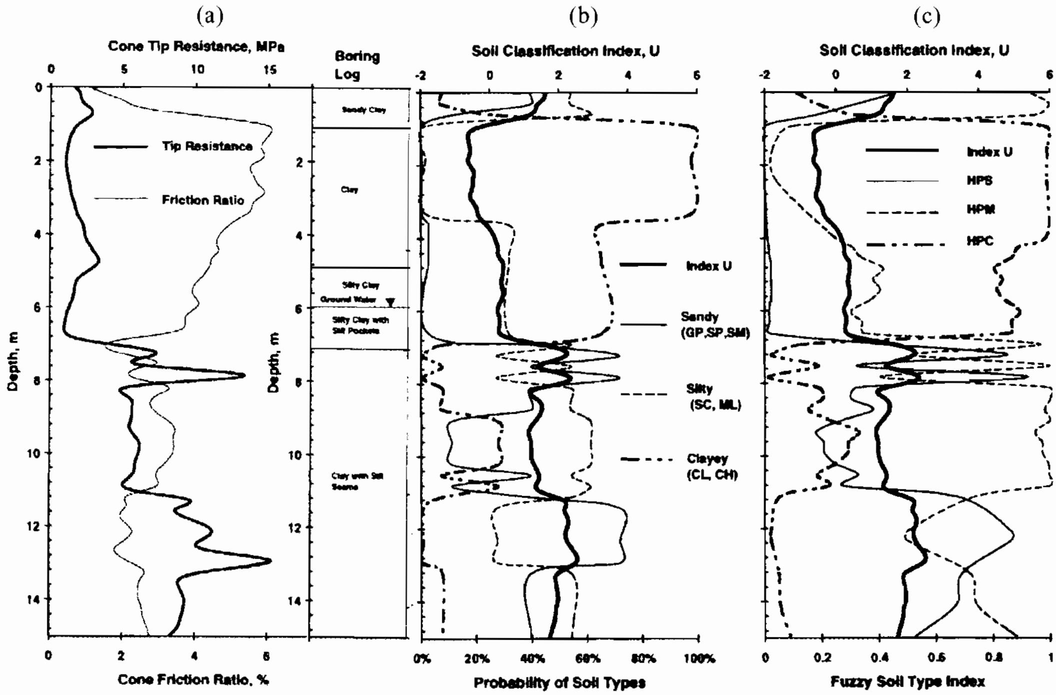

Figure 15. National Geotechnical Experimentation Site Texas A&M University: (a) CPT measurements; (b) probabilistic classification; (c) possibilistic classification (Zhang & Tumay 2003).

the likelihood (in probabilistic and possibilistic terms, respectively) that a soil whose mechanical properties have been measured belongs to a given soil type (in the compositional sense). The imperfect relation between soil composition and mechanical response to penetration was modelled statistically, through numerous comparisons of: (a) the value of a specifically defined soil classification index from couples of cone resistance and sleeve friction measurements; and (b) a compositional classification based on the USCS system. Subsequently, the value of such coefficient for each couple of measurements is associated with a measure of the degree of probability (or possibility) of being classified as one of the following three categories: (a) highly probable clayey soils [HPC], comprising USCS soil types GP, SP and SM; (b) highly probable mixed soils [HPM], comprising USCS soil types SC and ML; and (c) highly probable mixed soils [HPC], comprising USCS soil types CH and CL. The *membership functions* used in the possibilistic approach quantify the degree of membership (from 0: no membership to 1: full membership) of a given value of the soil classification index in the HPS, HPM and HPC fuzzy sets. For any set of CPT measurements, it is possible to describe continuous profiles in terms of soil behavior, quantifying, respectively, the *probability* and the *degree of possibility* that, at each depth, the penetrated soil behaves like a cohesionless, mixed or cohesive soil. Figure 15a reports the cone tip resistance and sleeve friction measurements as well as the descriptions pertaining to an adjacent boring. Figure 15b and Figure 15c illustrate the results of probabilistic and possibilistic classification of data. It may be observed that the probabilistic and possibilistic approaches provide information which is qualitatively consistent but quantitatively different. For instance, at each measurement depth the three probability values invariably add up to 100%; this is not a necessary condition for possibilistic classification.

9 ESTIMATION OF EPISTEMIC UNCERTAINTY

The procedures for the quantification of aleatory uncertainty, the assessment of statistical independence and stationarity, geostatistical kriging and the identification of homogeneous soil units reported in the previous sections implicitly neglect epistemic uncertainty, i.e. they assume that: (a) measured values reflect the true values of the parameter of interest; (b) the statistics calculated from samples are perfectly representative of the real populations; and (c) the geotechnical calculation models employed are perfect parameterizations of the relations between measurements and design values. None of these hypotheses is strictly acceptable.

The objective estimation of epistemic uncertainty is more difficult than that of aleatory uncertainty. While a variety of well established mathematical procedures (e.g. random field modeling, geostatistical kriging, wavelet theory, maximum likelihood methods) are available to pursue the estimation of inherent variability, a frequentist investigation on epistemic uncertainty requires data in sufficient number and of sufficient quality. Such conditions are very seldom met in routine geotechnical practice.

9.1 *Statistical estimation uncertainty*

Any quantitative geotechnical variability investigation must rely on sets (in statistical terms, *samples*) of measured data which are limited in size. *Sample statistics* are unable to represent the 'true' *population statistics* perfectly, in the sense that: (a) they may be biased; and (b) that there is some degree of uncertainty in their estimation. In other words, they are affected by *statistical estimation uncertainty*. If a data set consisting of n elements is assumed to be *statistically independent*, the expected value of the sample mean is equal to the (unknown) population mean; hence, the sample mean is an *unbiased estimator* of the population mean. However, the sample mean has a variance and, consequently, a standard deviation. The latter is given by $s_\xi\sqrt{n}$, in which s_ξ is the sample standard deviation and n is the numerosity of the data sample. The coefficient of variation of statistical estimation error is thus defined as

$$COV_{se} = \frac{s_\xi}{m_\xi\sqrt{n}} \tag{6}$$

Confidence intervals for sample statistics can also be calculated (see e.g. Ayyub & McCuen, 2003).

Statistical estimation uncertainty is epistemic in nature though its magnitude also depends on the COV of aleatory uncertainty of the random variable of interest. It can be effectively quantified using statistical theory, and generally reduced by increasing the number of available data though the standard deviation could increase with additional observations.

9.2 *Measurement uncertainty*

Measurement uncertainty results from a combination of several sources including the uncertainty associated with systematic testing error (i.e. equipment and operator/procedural effects), and random testing error which is not assignable to specific testing parameters. Measurement uncertainty can in principle be reduced by improving the quality of testing equipment, controlling testing conditions and ensuring strict adherence to test procedures.

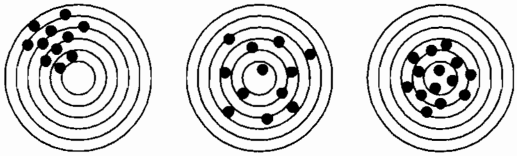

Figure 16. Target analogy for the conceptual distinction between accuracy and precision (adapted from Orchant et al. 1988).

Past research has made use of a variety of terms to qualify measurement uncertainty. Here, *accuracy* is defined as 'the closeness of agreement between any one measured value and an accepted reference value.' Orchant et al. (1988) defined *precision* as 'the closeness of agreement between randomly selected individual measurements or test results'. Hence, precision describes the repeatability of a test. The conceptual distinction between accuracy and precision may be appreciated in the 'target analogy' in Figure 16.

Measurement uncertainty is generally characterized statistically by a *measurement bias* (i.e. the possible consistent overestimation or underestimation of the real value of the object parameter, conceptually affine to accuracy), and a *measurement dispersion* (i.e. the scatter of measurements on presumably homogeneous soil volumes, conceptually related to precision). The above definitions, as well as results of past research focusing on the estimation of the magnitudes of the sources of laboratory and in-situ testing variability suggest that reliable direct quantification of measurement epistemic uncertainty requires repeated, comparative tests in replicate conditions (e.g. Orchant et al. 1988; Phoon & Kulhawy 1999a; 2005). Moreover, documentation on equipment and procedural controls during in situ testing is usually not detailed sufficiently to allow for a quantitative evaluation of measurement errors.

Literature values for design coefficients of variation of measurement uncertainty for a number of design parameters are shown in Table 4; coefficients of variation of measurement uncertainty for in-situ testing methods are given in Table 3. In some cases, values are assigned to a large extent subjectively on the basis of experience and judgment. In other cases, they are derived objectively by statistical estimation.

Neglecting measurement uncertainty is acceptable for tests which have been shown to be largely operator-independent and to have very low random measurement errors such as the CPTU or DMT (see Table 3). Results of tests with high measurement and random uncertainty (such as the SPT, which also has a

Table 3. Coefficients of variation of measurement uncertainty for in-situ testing methods (Kulhawy & Trautmann, 1996).

Test	Equipment	Oper./proc.	Random	Total[a]	Range[b]
Standard penetration test (SPT)	0.05[c]–0.75[d]	0.05[c]–0.75[d]	0.12–0.15	0.14[c]–1.00[d]	0.15–0.45
Mechanical cone penetration test (CPT)	0.05	0.10[e]–0.15[f]	0.10[e]–0.15[f]	0.15[e]–0.22[f]	0.15–0.25
Electric cone penetration test (ECPT)	0.03	0.05	0.05[e]–0.10[f]	0.07[e]–0.12[f]	0.05–0.15
Vane shear test (VST)	0.05	0.08	0.10	0.14	0.10–0.20
Dilatometer test (DMT)	0.05	0.05	0.08	0.11	0.05–0.15
Pressuremeter test, pre-bored (PMT)	0.05	0.12	0.10	0.16	0.10–0.20[g]
Self-boring pressuremeter test (SBPMT)	0.08	0.15	0.08	0.19	0.15–0.25[g]

[a] $COV(total)^2 = COV(equipment)^2 + COV(operator/procedure)^2 + COV(random)^2$.

[b] Because of statistical estimation uncertainty and subjective judgment involved in estimating COVs, ranges represent plausible magnitudes of measurement uncertainty for field tests.

[c,d] Best- to worst-case scenarios, respectively, for SPT.

[e,f] Tip and side resistances, respectively.

[g] It is likely that results may differ for p_0, p_f and p_L, but data are insufficient to clarify this issue.

Table 4. Literature values for design coefficients of variation of measurement uncertainty for design parameters (Phoon 2004; Lacasse, personal communication).

Property[a]	Test[b]	Soil type	Point COV	Spatial avg. COV[c]
s_u(UC)	Direct (lab)	Clay	0.20–0.55	0.10–0.40
s_u(UU)	Direct (lab)	Clay	0.10–0.35	0.07–0.25
s_u(CIUC)	Direct (lab)	Clay	0.20–0.45	0.10–0.30
s_u(field)	VST	Clay	0.15–0.50	0.15–0.50
s_u(UU)	q_t	Clay	0.30–0.40[e]	0.30–0.35[e]
s_u(CIUC)	q_t	Clay	0.35–0.50[e]	0.35–0.40[e]
s_u(UU)	N_{SPT}	Clay	0.40–0.60	0.40–0.55
s_u^d	K_D	Clay	0.30–0.55	0.30–0.55
s_u(field)	PI	Clay	0.30–0.55[e]	–
s_u(CAUC)	Direct (lab)	Clay	0.15–0.20	–
s_u(CAUE)	Direct (lab)	Clay	0.20–0.30	–
s_u(DSS)	Direct (lab)	Clay	0.10–0.15	–
ϕ'	Direct (lab)	Clay, sand	0.07–0.20	0.06–0.20
ϕ'(TC)	q_t	Sand	0.10–0.15[e]	0.10[e]
ϕ_{cv}	PI	Clay	0.15–0.20[e]	0.15–0.20[e]
K_0	Direct (SBPMT)	Clay	0.20–0.45	0.15–0.45
K_0	Direct (SBPMT)	Sand	0.25–0.55	0.20–0.55
K_0	K_D	Clay	0.35–0.50[e]	0.35–0.50[e]
K_0	N_{SPT}	Clay	0.40–0.75[e]	–
E_{PMT}	Direct (PMT)	Sand	0.20–0.70	0.15–0.70
E_D	Direct (DMT)	Sand	0.15–0.70	0.10–0.70
E_{PMT}	N_{SPT}	Clay	0.85–0.95	0.85–0.95
E_D	N_{SPT}	Silt	0.40–0.60	0.35–0.55

[a] s_u: undrained shear strength; UU: unconsolidated undrained triaxial compression test; UC: unconfined compression test; CIUC: consolidated isotropic undrained triaxial compression test; CAUC: consolidated anisotropic undrained triaxial compression test; CAUE: consolidated anisotropic undrained triaxial extension test; DSS: direct simple shear test; s_u(field): corrected su from vane shear test; ϕ': effective stress friction angle; TC: triaxial compression; ϕ_{cv}: constant volume effective friction angle; K_0: in-situ horizontal stress coefficient; E_{PMT}: pressuremeter modulus; E_D: dilatometer modulus.

[b] VST: vane shear test; q_t: corrected cone tip resistance from piezocone testing; N_{SPT}: standard penetration test blow count; K_D: dilatometer horizontal stress index; PI: plasticity index.

[c] Spatial averaging COV for an averaging distance of 5 m.

[d] Mixture of s_u from UU, UC and VST.

[e] COV is a function of the mean (see Phoon & Kulhawy 1999b).

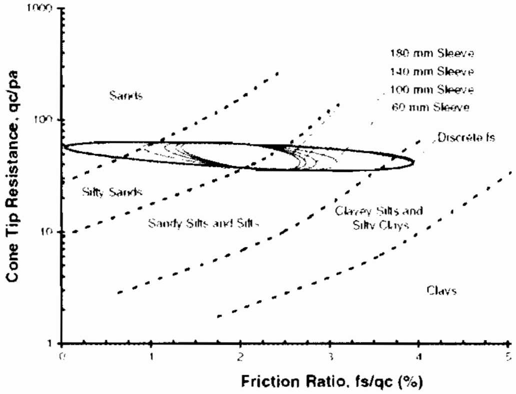

Figure 17. Effect of varying sleeve length on soil classification of simulated CPT data (Saussus et al. 2004).

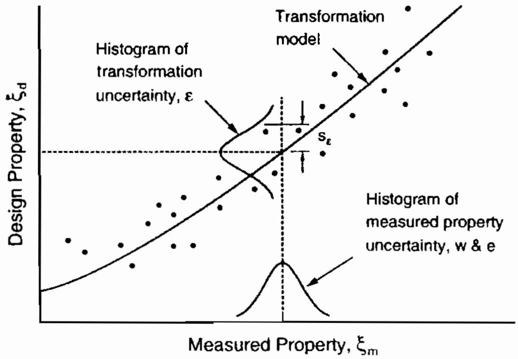

Figure 18. Probabilistic characterization of a geotechnical transformation model (Phoon & Kulhawy 1999b).

very large measurement interval, resulting in samples which are smaller in size and are thus affected by higher statistical estimation uncertainty) are most often not reliable inputs for such methods. The effects of measurement uncertainty can be assessed statistically. For instance, regarding the influence of measurement uncertainty on soil classification, Saussus et al. (2004) showed, using simulated profiles and analyzing the spatial correlation of CPT-related parameters, that CPT sleeve friction measurement introduces unnecessary redundancy due to the length of the standard friction sleeve compared to the measurement interval. Consequently, the smoothing and filtering of friction data affects, in the context of soil classification, the uncertainty in friction ratio values. Figure 17 shows an example result of the study, highlighting the significant effect of friction sleeve length on the classification of an artificially simulated sample realization of CPT measurements. It can be seen that the uncertainty in friction ratio increases with decreasing sleeve length. While studies of this type are not likely to be conducted on a routine basis, they are of extreme importance as they provide an idea of the relevant effect of testing equipment on important phases of geotechnical characterization and design.

9.3 *Evaluation of transformation uncertainty*

The direct measurement from a geotechnical test typically is not directly applicable for characterization or design purposes. Hence, a transformation (or calculation) model is required to relate test measurements to design parameters. In the process of model characterization, some degree of uncertainty is introduced. This is true whether the model is obtained by empirical fitting (because of inherent variability and epistemic uncertainty) or by theoretical analysis (because of the simplification of the physical reality).

An example of transformation model is the cone factor model, by which design undrained shear strength is estimated from CPT cone tip resistance (appropriately corrected in the case of piezocone data). The cone factor is calibrated for a specific site by regression analysis on CPT data and the undrained shear strength measured from laboratory tests. It should be highlighted, for this specific example, that due to shear strength anisotropy (different laboratory tests induce different types of failure in test samples) the cone factor is related to a specific test type, e.g. triaxial compression, triaxial extension, direct simple shear, etc.

Transformation uncertainty can be characterized statistically as shown in Figure 18. The transformation model is evaluated using regression analysis. In such an approach, the set of residuals of detrending are modeled as a zero-mean random variable. Transformation uncertainty is taken as the representative dispersion statistic (e.g. standard deviation or COV) of such random variable.

In practice, model characterization is not easily achieved, as reliable model statistics can only be evaluated by: (a) realistically large-scale prototype tests; (b) a sufficiently large and representative database; and (c) reasonably high-quality testing in which extraneous uncertainties are well controlled. Generally, available testing data are insufficient in quantity and quality to perform robust statistical assessment of model error in most geotechnical calculation models. Nonetheless, model uncertainty is of great importance in uncertainty-based design approaches such as reliability-based design, as its magnitude often exceeds that of other uncertainty components. Hence, efforts should be direct towards the compilation of a database of model statistics. Examples are given in Phoon & Kulhawy (1999b; 2005), Juang et al. (2005) and Uzielli et al. (2006b).

10 CLOSING REMARKS

Recognizing the existence of variability in geotechnical data for characterization and design purposes is the first step towards the consistent modeling, processing and assessment of the associated quantitative uncertainties. Statistical theory lies at the base of numerous methods and procedures which allow practical implementation of such uncertainty-based approaches by geotechnical engineers. The main benefits of such perspective for characterization and design are perhaps a more rational assessment of the degree of conservatism and risk and an improvement in the costs of investigation and design. This paper, far from aiming to be a self-standing contribution, aims to stimulate and accompany the growing diffusion of uncertainty-based approaches in geotechnical research and practice.

ACKNOWLEDGEMENTS

The author is grateful to the Norwegian Geotechnical Institute (NGI) for permitting use of the data for Figure 1, Figure 2 and Figure 7.

REFERENCES

Ang, A.H.S. & Tang, W.H. 1975. *Probability concepts in engineering planning and design.* New York: John Wiley & Sons.

Baecher, G.B. & Christian, J.T. 2003. *Reliability and statistics in geotechnical engineering.* New York: John Wiley & Sons.

Baise, L.G., Higgins, R.B. & Brankman, C.M. 2006. Liquefaction hazard mapping – statistical and spatial characterization of susceptible units. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE 132(6): 705–715.

Carr, J.R. 1995. *Numerical analysis for the geological sciences.* Englewood Cliffs: Prentice Hall.

Cherubini, C., Vessia, G. & Pula, W. 2006. Statistical soil characterization of Italian sites for reliability analyses. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the 2nd International Workshop on Characterisation and Engineering Properties of Natural Soils. Singapore, November 29–December 1, 2006.* The Netherlands: Taylor & Francis.

Chiasson, P. & Wang, Y.J. 2006. Spatial variability of sensitive Champlain sea clay and an application of stochastic slope stability analysis of a cut. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the 2nd International Workshop on Characterisation and Engineering Properties of Natural Soils. Singapore, November 29–December 1, 2006.* The Netherlands: Taylor & Francis.

Christian, J.T. 2004. Geotechnical engineering reliability: how well do we know what we are doing? *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE 130(10): 985–1003.

Dai, S.H. & Wang, M.O. 1992. *Reliability analysis in engineering applications.* New York: Van Nostrand Reinhold.

Daniel, W.W. 1990. *Applied nonparametric statistics – 2nd edition.* Boston: PWS-Kent.

Davis, J.C. 1986. *Statistics and data analysis in geology – 2nd edition.* New York: John Wiley & Sons.

Duncan, J.M. 2000. Factors of safety and reliability in geotechnical engineering. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE 126(4): 307–316.

El-Ramly, H., Morgenstern, N.R. & Cruden, D.M. 2002. Probabilistic slope stability analysis for practice. *Canadian Geotechnical Journal* 39: 665–683.

Facciorusso, J. & Uzielli, M. 2004. Stratigraphic profiling by cluster analysis and fuzzy soil classification. In A. Viana da Fonseca & P.W. Mayne (eds.), *Proceedings of the 2nd International Conference on Geotechnical Site Characterization ISC-2, Porto, 19–22 September 2004*: 905–912. Rotterdam: Millpress.

Harr, M.E. 1987. *Reliability-based design in civil engineering.* New York: McGraw-Hill.

Hegazy, Y.A. & Mayne, P.W. 2002. Objective site characterization using clustering of piezocone data. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 128(12): 986–996.

Hight, D.W. & Leroueil, S. 2003. Characterisation of soils for engineering practice. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Characterisation and engineering properties of natural soils*; *Proceedings of the first International Workshop on Characterisation and Engineering Properties of Natural Soils. Singapore, December 2–4, 2002*: 255–360. Lisse: Swets & Zeitlinger.

Jaksa, M.B. 1995. The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. *Ph.D. thesis*. University of Adelaide.

Jaksa, M.B. 2006. Modeling the natural variability of an overconsolidated clay in Adelaide, South Australia. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the 2nd International Workshop on Characterisation and Engineering Properties of Natural Soils. Singapore, November 29–December 1, 2006.* The Netherlands: Taylor & Francis.

Journel, A.G. & Huijbregts, C.J. 1978. *Mining geostatistics.* London: Academic Press.

Juang, C.H., Yang, S.H. & Yuan, H. 2005. Model uncertainty of shear wave velocity-based method for liquefaction potential evaluation. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE 131(10): 1274–1282.

Kulhawy, F.H. & Trautmann, C.H. 1996. Estimation of in-situ test uncertainty. In C.D. Shackleford, P.P. Nelson and M.J.S. Roth (eds.), *Uncertainty in the Geologic Environment: From Theory to Practice*, Geotechnical Special Publication No. 58: 269–286. New York: ASCE.

Lacasse, S. & Nadim, F. 1996. Uncertainties in characterising soil properties. In C.D. Shackleford, P.P. Nelson and M.J.S. Roth (eds.), *Uncertainty in the Geologic Environment: From Theory to Practice*, Geotechnical Special Publication No. 58: 49–75. New York: ASCE.

Lacasse, S., Guttormsen, T., Nadim, F., Rahim, A. & Lunne, T. 2007. Use of statistical methods for selecting design soil parameters. *Proceedings of the 6th International Conference on Offshore Site Investigation and Geotechnics, London, September 11–13, 2007.* London: Society for Underwater Technology.

Marchetti, S. & Crapps, D.K. 1981. Flat dilatometer manual. *Internal Report*. G.P.E. Inc.

Nadim, F. & Lacasse, S. 1999. Probabilistic slope stability evaluation. In *Proceedings of the 18th Annual Seminar on Geotechnical Risk Management, Hong Kong, May 14, 1989*: 177–186.

Orchant, C.J., Kulhawy, F.H. & Trautmann, C.H. 1988. Reliability-based foundation design for transmission line structures: critical evaluation of in-situ test methods. *Report EL-5507(2)*. Palo Alto: Electric Power Research Institute.

Phoon, K.K. & Kulhawy, F.W. 1999a. Characterisation of geotechnical variability. *Canadian Geotechnical Journal* 36: 612–624.

Phoon, K.K. & Kulhawy, F.W. 1999b. Evaluation of geotechnical property variability. *Canadian Geotechnical Journal* 36: 625–639.

Phoon, K.K. & Kulhawy, F.H. 2005. Characterisation of model uncertainties for laterally loaded rigid drilled shafts. *Géotechnique* 55(1): 45–54.

Phoon, K.K., Quek, S.T. & An, P. 2003c. Identification of statistically homogeneous soil layers using modified Bartlett statistics. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE 129(7): 649–659.

Popescu, R., Deodatis, G. & Nobahar, A. 2005. Effects of random heterogeneity of soil properties on bearing capacity. *Probabilistic Engineering Mechanics* 20: 324–341.

Réthâti, L. 1988. *Probabilistic solutions in geotechnics*. New York: Elsevier.

Robertson, P.K. 1990. Soil classification using the cone penetration test. *Canadian Geotechnical Journal* 27: 151–158.

Robertson, P.K. & Wride, C.E. 1998. Evaluating cyclic liquefaction potential using the cone penetration test. *Canadian Geotechnical Journal* 35: 442–459.

Snedecor, G.W. & Cochran, W.C. 1989. *Statistical methods*. Ames: University of Iowa Press.

Uzielli, M. 2004. Variability of stress-normalized CPT parameters and application to seismic liquefaction initiation analysis. *Ph.D. thesis*. University of Florence, Italy.

Uzielli, M., Vannucchi, G. & Phoon, K.K. 2004. Assessment of weak stationarity using normalised cone tip resistance. *Proceedings of the ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, Albuquerque, New Mexico, July 26–28, 2004* (CD ROM).

Uzielli, M., Vannucchi, G. & Phoon, K.K. 2005a. Random field characterisation of stress-normalised cone penetration testing parameters. *Géotechnique* 55(1): 3–20.

Uzielli, M., Vannucchi, G. & Phoon, K.K. 2005b. Investigation of correlation structures and weak stationarity using the CPT soil behaviour classification index. In G. Augusti, G.I. Schuëller & M. Ciampoli (eds.), *Safety and Reliability of Engineering Systems and Structures – Proceedings of ICOSSAR 2005, Rome, 19–23 June, 2005*. Rotterdam: Millpress (CD ROM).

Uzielli, M., Lacasse, S., Nadim, F. & Phoon, K.K. 2006a. Soil variability analysis for geotechnical practice. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the Second International Workshop on Characterisation and Engineering Properties of Natural Soils*. Singapore, November 29–December 1, 2006. The Netherlands: Taylor & Francis.

Uzielli, M., Lacasse, S., Nadim, F. & Lunne, T. 2006b. Uncertainty-based analysis of Troll marine clay. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the 2nd International Workshop on Characterisation and Engineering Properties of Natural Soils. Singapore, November 29–December 1, 2006*. The Netherlands: Taylor & Francis.

Uzielli, M. & Mayne, P.W. 2008. Comparative CPT-based classification of Cooper Marl. *Proceedings of the 3rd International Conference on Site Characterization, Taiwan, April 1–4, 2008*. The Netherlands: Taylor & Francis.

Uzielli, M., Simonini, P. & Cola, S. 2008. Statistical identification of homogeneous soil units for Venice lagoon soils. *Proceedings of the 3rd International Conference on Site Characterization, Taiwan, April 1–4, 2008*. The Netherlands: Taylor & Francis.

Vanmarcke, E.H. 1983. *Random Fields: analysis and synthesis*. Cambridge: MIT Press.

Zhang, Z. & Tumay, M.T. 2003. Non-traditional approaches in soil classification derived from the cone penetration test. In G.A. Fenton & E.H. Vanmarcke (eds.), *Probabilistic Site Characterisation at the National Geotechnical Experimentation Sites*, Geotechnical Special Publication No. 121: 101–149. Reston: ASCE.